Does Teacher Experience Matter? Comparing the Effect of Class Size and Teacher Experience on Educational Outcomes

Utkarsh Dandanayak

November 20th, 2022

1 Introduction

There exists an extensive literature which considers the effects of reducing class sizes on students' academic performance, perhaps due to its popularity as an educational policy. Indeed, "a common response to overcrowded classrooms is to add more teachers" (The Royal Swedish Academy of Sciences 2019). By contrast, there is significantly less research available that considers the effect of teacher experience on students' academic achievement. It is important, however, that such a relationship be examined – it has been demonstrated in the past that policies pertaining to the **type** of teacher employed in the classroom can be significantly more effective than simply reducing the student-teacher ratio. For example, one study in Kenya found that reducing class sizes alone "generated little improvement in test scores", but when students were taught by newly-hired contract teachers (as opposed to the existing civil-service teachers) in these smaller classes, students did perform better in tests (Duflo, Dupas, and Kremer 2012). In this paper, I examine the relationship between class size and student test scores, and teachers' years of experience and student test scores, in order to determine which of the two models provides a stronger indicator for academic performance. This will contribute to existing literature by investigating an otherwise understudied relationship between academic performance and teacher experience, and comparing it to a conventional policy proposal that **has** been extensively studied (i.e. reducing class sizes). This would allow policymakers to make more well-informed decisions on education policy – this study would enable them to determine whether hiring younger, more inexperienced teachers is more cost effective than retaining an older, more experienced workforce (which may be more costly to maintain, due to the higher wages generally paid to workers with more experience).

2 Methodology

I will be using a two-variable linear regression in order to model the relationship between test scores/class size and test scores/teacher experience, using data for these variables in the *star.dta* dataset, which contains 5710 observations. I am concerned with how changing class sizes and teacher experience would affect students' **holistic** academic performance, rather than performance in a specific subject or test. For this reason, I created a new variable, **total_score**, which sums each student's scores in math, listening, word study skills, and reading tests. Then, I regressed **total_score** against **teacher_experience** (which records each teacher's total number of years of teaching experience), and also regressed **total_score** against **class_size** (which

records the total number of students in each student's class). In other words, **total_score** is the dependent variable, while **teacher_experience** and **class_size** are the independent variables in our regressions.

3 Results

The results from each regression are displayed below:

Source	SS	df	MS	Number of obs	=	5,710
				F(1, 5708)	=	86.31
Model	1434805.92	1	1434805.92	Prob > F	=	0.0000
Residual	94888178.9	5,708	16623.7174	R-squared	=	0.0149
				Adj R-squared	=	0.0147
Total	96322984.8	5,709	16872.1291	Root MSE	=	128.93

total_score	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
teacher_experience	2.751549	.2961725	9.29	0.000	2.170939	3.33216
_cons	1869.093	3.246646	575.70	0.000	1862.728	1875.458

Figure 1: Results from regression of student test scores against teachers' years of experience

Source	SS	df	MS	Number of obs	=	5,710
				F(1, 5708)	=	42.67
Model	714730.105	1	714730.105	Prob > F	=	0.0000
Residual	95608254.7	5,708	16749.8694	R-squared	=	0.0074
				Adj R-squared	=	0.0072
Total	96322984.8	5,709	16872.1291	Root MSE	=	129.42

total_score	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
class_size	-2.82135	.4319084	-6.53	0.000	-3.668054	-1.974645
_cons	1951.977	8.925813	218.69	0.000	1934.479	1969.475

Figure 2: Results from regression of student test scores against class size

The coefficients for the regression line of best fit are approximately 2.75 (teacher experience) and -2.82 (class size). We can interpret these coefficients as follows - a 1 year increase in the teacher's years of experience corresponds with a 2.75 point increase in their students' total test scores, and a 1 person decrease in a student's class size corresponds with a 2.82 point increase in that student's total test scores.

Note that the p-values are 0.000 for both regressions, implying that both regression results are statistically significant even at the 0.1% level ($p < 0.001$). However, one important distinction between the regressions is the difference in R-squared values. The R-squared value for the teacher experience regression is 0.0149, meaning that teacher experience explains 1.49% of the variability in student test scores. The R-squared value for the class size regression is even lower, at 0.0074, meaning that class size can only explain 0.74% of

the variability in student test score data. Thus, the data more closely fits the regression model which uses teacher experience, rather than class size, as the independent variable, suggesting that teacher experience is a stronger indicator for test scores compared to class sizes. However, these relatively small R-squared values imply that neither variable is a particularly effective indicator for student test scores. Each line of best fit can be visualized below, on the same axes as a scatter plot of the data itself:

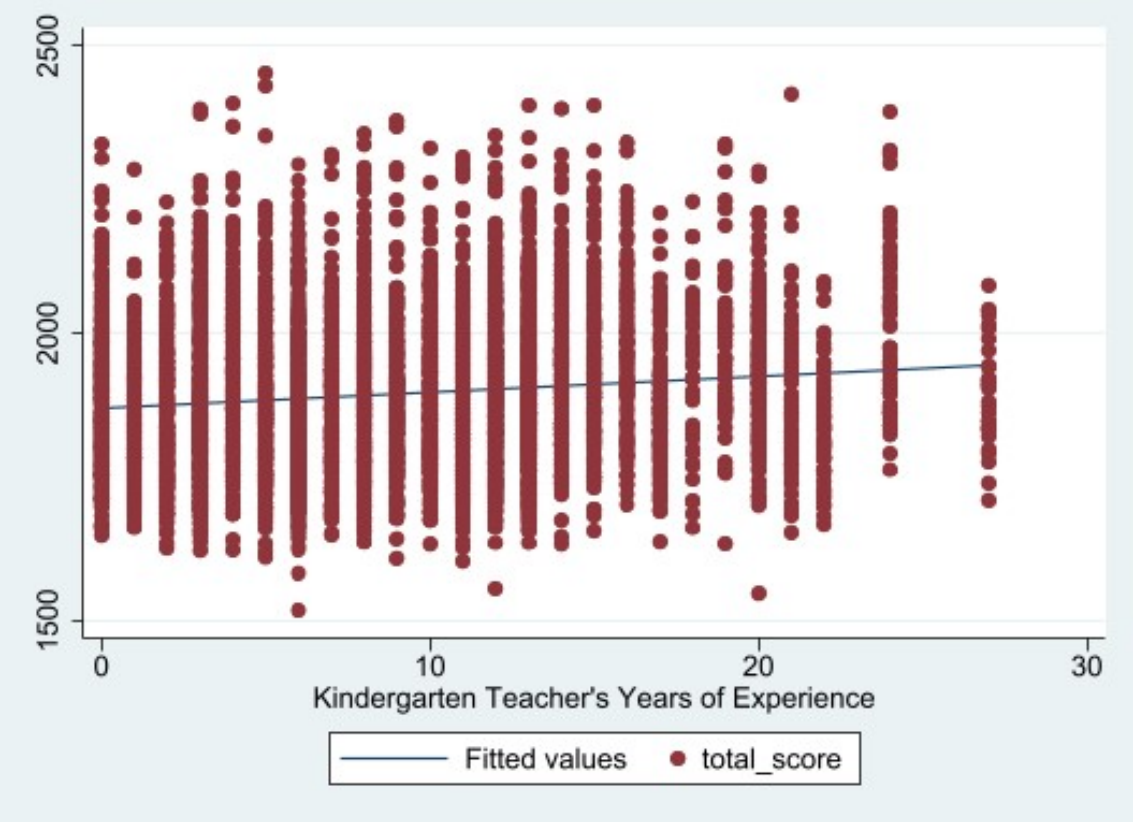

Figure 3: Scatter plot and line of best fit for student test scores against teachers' years of experience

4 Conclusion

In summary, the above analysis reveals that although there is likely to be a non-zero correlation between teacher experience and test scores, and class size and test scores (since $p < 0.001$), both models are not particularly strong in their ability to explain the student test score data. While the data did fit the teacher experience model slightly better, both R-squared values were very low, which means that both teacher experience and class size are weak indicators of students' overall academic achievement. This suggests that economic policies focussed on maintaining a more experienced workforce (e.g. restricting the hire of new entrants into the teaching workforce, and instead hiring experienced teachers) are unlikely to translate into a tangible positive outcome for student test scores. Similarly, economic policies which encourage hiring teachers to reduce student-teacher ratios are also unlikely to be effective in improving students' educational outcomes (the latter finding agrees with the conclusions drawn in Duflo, Dupas, and Kremer 2012). It may still be beneficial to construct a more concrete analysis of the potential causal relationship between teacher

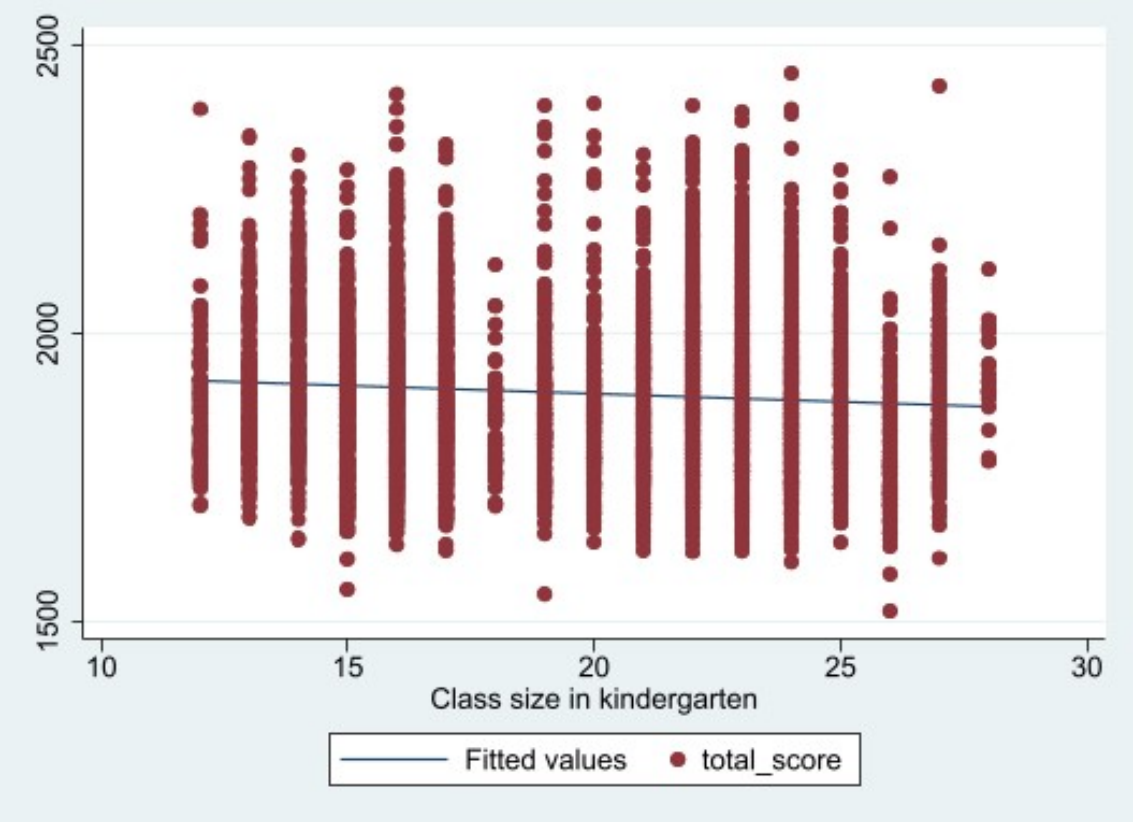

Figure 4: Scatter plot and line of best fit for student test scores against class size

experience and test scores - this could be achieved with a randomized controlled trial (RCT) that randomly assigns students to classes run by teachers with more experience and teachers with less experience. This experimental setup would enable a researcher to control for confounding variables and make some ceteris paribus assumptions, in order to isolate the causal interaction between student test scores and teacher experience.

References

[DDK12] Esther Duflo, Pascaline Dupas, and Michael Kremer. "School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools". In: *SSRN Electronic Journal* (Mar. 2012). DOI: 10.2139/ssrn.2021483. (Visited on 11/20/2022).

[Sci19] The Royal Swedish Academy of Sciences. *Understanding Development and Poverty Alleviation*. Oct. 2019. URL: https://www.nobelprize.org/uploads/2019/10/advanced-economicsciencesprize2019.pdf.